



# HHS Public Access

Author manuscript

*Med Care*. Author manuscript; available in PMC 2019 October 01.

Published in final edited form as:

*Med Care*. 2018 October ; 56(Suppl 10 1): S64–S69. doi:10.1097/MLR.0000000000000786.

## A Tale of Two Constituencies: Exploring Patient and Clinician Perspectives in the Age of Big Data

Crispin N Goytia, BS<sup>1</sup>, Isaac Kastenbaum, MPA<sup>2</sup>, Donna Shelley, MD, MPH<sup>3</sup>, Rainu Kaushal, MD, MPH<sup>4</sup>, and Carol R. Horowitz, MD, MPH<sup>1</sup>

<sup>1</sup>Icahn School of Medicine at Mount Sinai, NYC, NY

<sup>2</sup>New York Presbyterian Hospital, NYC, NY

<sup>3</sup>New York University School of Medicine, NYC, NY

<sup>4</sup>Weill Cornell Medical College, NYC, NY

### Abstract

**Background**—Patient and clinician stakeholders are inadequately engaged in key aspects of research, particularly regarding use of Big Data to study and improve patient-centered outcomes. Little is known about the attitudes, interests and concerns of stakeholders regarding such data.

**Research Design**—The NYC Clinical Data Research Network (CDRN), a collaboration of research, clinical and community leaders built a de-identified dataset containing electronic health records from millions of New Yorkers. Guided by a patient-clinician advisory board, we developed a question guide to explore patient and clinician experiences and ideas about research using large datasets. Trained facilitators led discussions during pre-existing patient, community and clinician group meetings. The research team coded meeting notes and identified themes.

**Results**—Fully 272 individuals participated in 19 listening sessions (139 patients/advocates, 133 clinicians) at 6 medical centers with diverse NYC communities: 76% were female and 63% were non-white. Clinicians and patients agreed on all major themes including the central role of clinicians in introducing patients to research and the need for public campaigns to inform stakeholders about Big Data. Stakeholders were interested in using granular data to compare the

---

**Author Correspondence Information:** Crispin Noelle Goytia, BA, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, 1 Gustave Levy Place, Box1077, New York, New York 10029, Telephone: 212-824-7323, Fax: 212-423-2998, [crispin.goytia@m Mountsinai.org](mailto:crispin.goytia@m Mountsinai.org).  
Crispin Noelle Goytia, BA, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, 1 Gustave Levy Place, Box1077, New York, New York 10029, Telephone: 212-824-7323, Fax: 212-423-2998, [crispin.goytia@m Mountsinai.org](mailto:crispin.goytia@m Mountsinai.org)  
Isaac Kastenbaum, MPA, New York-Presbyterian Hospital, 525 E 68<sup>th</sup> Street, Mailbox 259, NY, NY 10065, Telephone: 484-794-1768, Fax Number: 212-746-4899, [ink9012@nyp.org](mailto:ink9012@nyp.org)  
Donna Shelley, MD MPH, New York University School of Medicine, Department of Population Health, 227 east 30th street, New York, NY 10016, Telephone: 6465013526, Fax: 6465012706, [Donna.Shelley@nyumc.org](mailto:Donna.Shelley@nyumc.org)  
Carol R. Horowitz, MD MPH, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, 1 Gustave Levy Place, Box1077, New York, New York 10029, Telephone: 212659-9567, Fax: 212-423-2998, [Carol.Horowitz@mssm.edu](mailto:Carol.Horowitz@mssm.edu)  
Rainu Kaushal MD, MPH, Weill Cornell Medicine, New York-Presbyterian, Department of Healthcare Policy & Research, 402 East 67<sup>th</sup> Street, Floor 2, New York, NY 10065, Telephone: 646-962-8009, Fax: 646-962-0281, [rak2007@med.cornell.edu](mailto:rak2007@med.cornell.edu)

### Conflict of Interest:

The authors do not have any conflict of interest to report.

care and clinical outcomes of their neighborhoods with others across NYC, but were also concerned that data could not truly be de-identified.

**Conclusion**—Clinicians and patients agree on potential benefits of stakeholder-engaged Big Data research and provided suggestions for further research and building stakeholder research capacity. This evaluation demonstrated the potential of brief meetings with existing patient and clinical groups to explore barriers and facilitators to patient and clinician engagement.

### Keywords

patient engagement; clinician engagement; Big Data; patient-centered outcomes

---

## Introduction

Promotion and support of stakeholder involvement in research can influence healthcare decisions, organizational design, governance, policy making and health outcomes.<sup>12</sup> Actively engaged patients, caregivers, advocates, and clinicians can identify novel questions and approaches, enhance the success of research initiatives and the translation of findings into sustainable practices.<sup>345</sup> However, diverse patients and busy practicing clinicians have not been adequately engaged in study design, implementation and analysis.<sup>6789</sup> There is a growing body of data on barriers to engagement of these stakeholders, even in research that may intuitively be more understandable and attractive, e.g., comparative effectiveness and community engaged research. Barriers include competing demands for time and attention, research mistrust, and inadequate understanding of the promises of research, especially among individuals from underserved populations, and those without previous research exposure.<sup>10111213</sup>

While roadmaps to appropriately engage stakeholders in research are emerging,<sup>14</sup> there is very limited understanding of how to engage stakeholders in the rapidly evolving and increasingly complex field of Big Data (data science). Big Data efforts, including the Patient-Centered Clinical Research Network (PCORnet) funded by the Patient-Centered Outcomes Research Institute,<sup>151617</sup> the National Institutes of Health's Big Data to Knowledge program<sup>18</sup> and the Precision Medicine Initiative<sup>19</sup> aim to gather data from multiple sources to study and improve healthcare and health. These large national initiatives are beginning to study whether and how stakeholders understand the concepts, uses, promises and perils of Big Data, are prepared to contribute to and shape Big Data initiatives, where their interests lie, and to begin to build networks or communities of engaged stakeholders.

The NYC Clinical Data Research Network (NYC-CDRN), a member of PCORnet, brings together over 20 organizations, including 6 independent health systems contributing data to a central de-identified data repository. Its collaborative governance model leverages members' expertise and experience to inform patient-centered, comparative effectiveness research.<sup>20</sup> The NYC-CDRN focuses on socioeconomically, racially and ethnically diverse patient populations and clinicians in a geographically defined urban area. Like other members of Big Data initiatives, our seasoned researchers, clinicians, advocates, technology, privacy and security experts rapidly recognized the lack of an established framework for patient and

clinician engagement in Big Data research. To inform our engagement strategies and prioritize research topics, we aimed to gain insight from stakeholders into their understanding of Big Data, interest and concerns in contributing to this research and questions they would like to answer.

## Methods

The NYC-CDRN's organizational structure includes a Governance Board, and four subcommittees: patient/clinician engagement, privacy and security, technology, and comparative effectiveness/research. Engagement subcommittee leaders conceived of and led this study, but worked closely with the other subcommittees, the governance board, and a newly-formed 16-member Patient-Clinician Advisory Board on all aspects of the study. First, leaders held discussions with each subcommittee to gather the questions they would like to ask of patients, advocates and clinicians in order to accomplish their aims. These discussions also allowed us to introduce, in detail, the purpose and benefits of stakeholder engagement. Second, we reviewed these questions with the Advisory Board, who added additional questions and worked with us to develop a recruitment strategy.

Rather than conducting formal 60–90-minute focus groups, we decided to conduct “opportunistic” listening sessions, asking clinician, patient and community advocacy groups who were identified by the Advisory Board. These groups have regularly scheduled, well-attended meetings, to allow us to join their meetings and engage their members in discussions for up to 30 minutes. The clinician groups consisted of faculty/staff meetings. The patient/advocacy groups consisted of advocacy groups, social service organizations, patients and disease groups (rare and chronic). The rationale was to garner input from more representative stakeholders, including those who may not attend a separate focus group, such as the busiest stakeholders, those with limited trust and initial interest in Big Data or research. By answering open-ended questions, the group could think about this new area, and share ideas that do not emerge from surveys with their pre-populated answer choices for pre-selected questions.

In partnership and with approval of the Patient/Clinician Advisory Board and the Governance Board, the subcommittee developed moderator guides including questions about previous participation in research, attitudes, pros and cons about Big Data research and the NYC-CDRN. The questions included: what they would want to ask using the data, ways they would share information about the CDRN, how to garner patient and clinician participation in CDRN and ways to reach out to patients in the CDRN to solicit their participation in future studies. We also developed, a one-page introduction for attendees of listening sessions, describing the NYC-CDRN and Big Data research. The Advisory Board worked with the privacy committee to address concerns regarding consent. They both agreed consent was not necessary given that no identifying data from participants was being collected. Also, the study was deemed exempt from informed consent by the IRB. Those who chose not to participate were given the opportunity to leave the session. Participants were notified that any results from the session would be used in publication(s) and disseminated to all groups who participated.

To recruit groups for listening sessions, we worked with members of all CDRN subcommittees, and the Advisory and Governance Boards, to identify pre-existing groups and ascertain their interest in taking part in listening sessions. We aimed to recruit geographically, demographically and clinically diverse groups. Because the initial focus of our CDRN was on two common, chronic conditions (diabetes, obesity) and one rarer chronic condition (cystic fibrosis), we tailored recruitment to reach relevant groups. All groups solicited chose to participate.

Trained qualitative researchers led each session, and trained note-takers recorded key elements of each discussion and basic demographics of participants (gender, geographic location, race/ethnicity, and for clinicians, their specialty). We began with a brief discussion of the NYC-CDRN, then led participants through the questionnaire guide, asked for feedback on our one-pager, and ascertained their interest in being involved with future CDRN-activities.

We analyzed discussion notes and field notes taken by the researchers using thematic analysis,<sup>21,22</sup> and two independent coders assigned codes to the notes and compared codes between patient/advocate and clinician groups. We then reviewed the codes to develop themes unique to each group (patient/advocate and clinician) and common to both, used member checking to ascertain validity,<sup>23</sup> and calculated inter-rater reliability (with kappa statistic of 92%).

## Results

There were 19 listening sessions. The 11 clinician groups had 133 (range 10–30/group) academic and community physicians from cardiology, endocrinology, family medicine, internal medicine, nephrology, pediatrics and pulmonary, as well as nutritionist and nurses. The 8 patient/advocate groups had 139 participants (range 8–20 people/group) from free standing community organizations and disease support groups from various NYC neighborhoods with disease including diabetes, obesity, cystic fibrosis, cardiovascular disease, hypertension and depression. Most (63%) were non-white (91% of patients, 44% of clinicians) and 76% were female (85% of patients, 67% of clinicians). Few reported experience as participants in research studies (33% of patients; 0% of clinicians).

Five major themes emerged. All are common to patient and clinician groups: (Tables 1–4) (1) Clinicians should be central to the process of engaging patients in research; (2) There is a need to build capacity for stakeholders to translate their thoughts and ideas into research questions; (3) There are primarily interested in using data to compare different health centers/hospitals, or different neighborhoods characteristics and outcomes to learn how and why their site/community differs from others; (4) They recommend a city-wide, or even nation-wide campaign to explain to the public why Big Data matters; and (5) While Big Data studies holds promise, there are concerns about privacy and transparency.

### Theme 1: Clinicians are important to engage patients in research

The first theme is that clinicians should be central to patient engagement in research, particularly because benefits of research are not readily apparent. Clinicians should consider

the merits of studies, grant access to their patients, and inform patients about research opportunities. In large part this is because many participants believe that research in general is a good idea, but something is “lost in the translation” in engaging patients and clinicians in research. One patient said, “Researchers are in outer space from what patients are going through.” Patients and clinicians did not appreciate a significant impact of much of research done to date. Clinicians discussed how findings and guidelines keep changing and studies sometimes offer little new actionable information. Thus it may be wiser to spend the limited time they have with patients doing things they know they need to do, than to spend time on research, with unproven benefits. They are also concerned that patients from diverse backgrounds, with limited literacy or formal education will not respond well to researchers, will refuse to participate in important studies if not approached by their clinicians first, and therefore research findings will not be meaningful to these populations. Patients noted that they cannot readily think of examples of how their lives and health improved based on research, making it less attractive. And, as they did not separate their attitudes toward clinical care and research, if they have a negative clinical experience, they assume they will have a negative research experience. These concerns could be mitigated for patients who trust their clinicians, by receiving a “warm handoff” from their clinician to the research staff. Clinicians would be more comfortable if patients are not asked to participate in a research study without their knowledge, and if they first have time to understand the research and its potential patient burden, risk and impact. Both groups also noted that positive past research experiences would foster positive future experiences, and vice versa.

### **Theme 2: Build Capacity for stakeholders to translate ideas into research questions**

The second theme is that both groups need assistance in learning how to translate their concerns and ideas into concrete research questions. Big Data research is a relatively new concept to many, and when asked what kinds of questions they would like to ask using Big Data, they stated they did not know how to convert the many ideas they had into solid research questions that would pique the interest of researchers. Many stated that with appropriate effort, researchers could demystify research, help people develop ideas and thereby create positive research experiences which would facilitate future engagement efforts.<sup>24</sup> Clinicians are worried their patients will refuse to participate in a study due to a lack of research literacy, particularly in the new world of Big Data.

### **Theme 3: Use data to compare different health centers/hospitals/neighborhood characteristics and outcomes**

The third theme was that when asked to provide ideas for research, nearly all ideas involved comparing how their clinic or neighborhood compared with other clinics or neighborhoods. How do care and outcomes differ for a given condition? Are patients in some neighborhoods sicker than in others, and is this due to clinical or neighborhood factors? Do some clinical sites offer better treatments for a given disease that lead to better outcomes? Rather than groups honing in on specific research questions, they honed in on conducting comparison studies.

**Theme 4: Recommend a city/nation wide campaign to explain importance of research**

The fourth theme shared by patients and clinicians was the need for larger campaigns to introduce their stakeholder groups to Big Data research. Many patients are unaware that clinical information ever goes beyond their doctors' offices to places such as insurers or regulators. As the introduction to the NYC-CDRN was the first time they learned of this, they had an initial negative reaction to the CDRN's plans. One patient said, "Let me know what hospitals do not share their data with anyone, so that I can go do a doctor at one of them." Interestingly Kim et.al, discuss in their study that trust is still a very real issue for people and it can affect a person's perception of who they should allow access to their healthcare information.<sup>25</sup> However, as they learned how data is safely and securely shared, many patients and clinicians warned that people like them may have a similar negative reaction if they are not first informed about how their data is currently used outside clinical settings. They think it is too much to ask the NYC-CDRN to educate all of NYC on why this matters, and suggest that national agencies develop educational campaigns. They also suggest that having patient and clinician-friendly portals to the data will help people see the promise of such endeavors (particularly if they can be used to compare their group to others, as in theme three), and help gain wide support for Big Data research. Both groups also suggest having Big Data portals be linked to information about the diseases being studied and how diagnoses and treatments are informed by past research, to help patients understand how research can change lives.

**Theme 5: Big Data studies hold promise, but concerns about privacy and transparency**

The fifth theme is that while recognizing some of the promise of Big Data, both groups express concerns about potential for breaches of privacy. This is similar to Ancker's Cornell Study where survey participants cited privacy and security of their data as serious concern.<sup>26</sup> Clinicians are also concerned that the data collected will not be accurate (i.e., the problem lists and demographics may have inaccuracies), will lead to inaccurate findings and inappropriately impact treatment guidelines. Similarly, patients/advocates state that "bad" data can have negative implications and create stigma for communities, especially among low socio-economic populations. Both groups also state that the findings from research may be kept within ivory towers, not benefitting front-line clinicians or patients from diverse communities. In addition, patients are worried data can never be truly de-identified and other hospitals or parties will have access to their data without their knowledge.

Interestingly, there is only one area where patients and clinicians differed, namely how to reach out to patients to solicit research participation. Generally, patients stated they prefer to be contacted through regular postal mail and phone calls from their clinicians or sites of care. Many were uncomfortable with the idea of receiving emails or texts from strangers. In contrast, clinicians suggested engaging patients by text message and email, even expressing disagreement when facilitators mentioned that patients had concerns about this. Clinicians believed that, with the increase use of technology, research staff will have more success engaging patients in research.



## Discussion

Stakeholder engagement is a crucial piece of the research puzzle and is largely unexplored in the area of Big Data. Before embarking on comparative effectiveness research using a centralized, de-identified large clinical dataset, NYC-CDRN leaders conducted listening sessions with patients, advocates and clinicians to understand their concerns, ideas and suggestions. Themes uncovered were held in common by both clinicians and patients.

Both groups prefer clinicians have a central role in engaging their patients in research, as trusted judges of good research and protectors of their patients. Others have suggested it is neither necessary nor advisable for clinicians to have this role, and patients can and should decide for themselves.<sup>27</sup> Interestingly, patients do not seem to distinguish between research and clinical care from the same site, such that positive clinical experiences may facilitate participation, but negative clinical experiences may thwart participation. To overcome research inertia- not having time or not seeing sufficient benefit to taking part in research, particularly in the new and harder to understand world of Big Data, stakeholders believe clinicians hold a very important role.

Stakeholders also want more information to help them understand Big Data and formulate their own questions, because they do not know if research already exists for what they are interested in. The populations most impacted by this research-diverse patient and the clinicians who care for them, have the best understanding of what the challenges and potential solutions are to improve health, so it is crucial to facilitate their input. Clinicians and patients may require assistance with clarifying their clinical focus and translating ideas into a research question.<sup>28</sup> Several organizations, including ours, are building programs to increase capacity for this input into big data analytics, and ongoing research will determine which approaches are most effective.<sup>29</sup>

Patients and clinicians suggested that research leaders explain that clinical data is used outside of doctor's offices for many purposes. CDRN-like organizations aim to distinguish themselves by ensuring that patient and clinicians are at the table to use these data wisely and effectively. However, if stakeholders hear, for the first time, that this data is not held solely by their clinicians, their negative reaction to sharing data may be singularly focused on groups who are simply making the most effort to make sure stakeholders understand when and how their data is used. Patients do have privacy concerns with big data analytics.<sup>30</sup> Our patient partners suggested that larger entities introduce the concept of clinical data warehouses, so such information will not cause a backlash and damn our efforts to engage the stakeholders. In the New York Times, Steve Lohr said, "But the latest leaps in data collection are raising new concern about infringements on privacy — an issue so crucial that it could trump all others and upset the Big Data bandwagon."<sup>31</sup> It may be complicated to have people understand that their data is "out there" already, react to that, possibly negatively, and then follow on with supplemental point that this time, they can use it.<sup>32</sup> More research is needed on how to develop a systematic framework to open dialogues with patients about how to explain the use of clinical data in general, and for patient-centered research in particular.<sup>33</sup>

Our team found it interesting that on one hand, patients and clinicians wanted the very granular data, namely information about site of medical care and neighborhood patients live in, in order to compare risks, processes and outcomes. Indeed, such groups were asking that data on social determinants of health be included in these explorations, much of which is only accurate in large cities, at block levels. On the other hand, they are concerns that data cannot be truly de-identified, and this type of granular data makes de-identification challenging. The conflict between protecting privacy and answering patient and clinician-centered questions deserves further exploration. Privacy concerns can inhibit stakeholder engagement; however, transparency is necessary for trust and growth to flourish. It is imperative to be transparent with your stakeholders throughout the entire process. Some patient populations are increasingly interested in sharing their data to advance health, particularly for conditions they have, and feel that direct benefits from data they contribute outweigh protections of restrictive rules that may thwart research.<sup>34</sup> However, our stakeholders did not seem that enthusiastic.

Patients and clinicians also express concerns about the accuracy of clinical data. If data used is not accurate, they suggest it could stigmatize patients, or yield inadequate data on some populations (for example by not accurately capturing racial, ethnic, language or sexual minority groups), give doctors and their practices bad reputations and lead to inaccurate conclusions that could harm clinical care. Procedures leading to inaccurate or biased data may be damaging and are hard to detect.<sup>35</sup> Given these concerns; perhaps stakeholders could become allies in efforts to improve the accuracy and completeness of electronic health records.

Our study does have limitations. These were half hour group interviews, and while we took extensive notes, we did not audiotape and transcribe the sessions. The listening sessions were conducted in only one urban area and confined to the NYC-CDRN. However, this form of “quick and dirty” data gathering may have some advantages that deserve consideration. First, they may be an efficient way to obtain information from groups who will not have the time or inclination to join formal focus groups, and do so in a setting and among people they are comfortable with. Participants have stated the major reason for lack of participating in focus groups was the time burden.<sup>36</sup> This only reaffirms why listening sessions are a more efficient strategy to engage stakeholders.

We were also able to engage with diverse groups, garnering a wider spectrum of ideas, and from a larger network of individuals. In addition, all groups we met with agreed enthusiastically that we could return with further questions and to field future ideas. In this way, we may have created a network of networks: stakeholders who represent other stakeholders, with whom we can work, thereby garnering input way beyond the confines of the important, but rather small stakeholder boards traditionally involved in engaged research. We will disseminate to both clinician and patient/advocacy groups to facilitate knowledge gaps concerning research and large datasets. Also, we hope to gain feedback on the process and themes from these groups. Others working in research may want to consider this as a viable approach to garnering patient and clinician perspectives.



More research is needed to determine best ways to engage stakeholders in Big Data research. We did not find similar studies that addressed patient engagement and big data in the manner we chose. Our results may be generalizable to other CDRNs interested in patient/clinician engagement frameworks in Big Data. We found great agreement between clinicians and patients about how to engage and collaborate in research and are beginning to chart a road forward, forming network of networks. CDRNs can use these themes to create a research capacity building program for patients and clinicians and integrate their perspectives into research studies. The literature suggests that utilizing existing social networks can expand research knowledge and interest among stakeholders as well as create opportunities to engage diverse populations in research.<sup>37</sup>

## Acknowledgments

### Disclosure of Funding Sources:

This work was funded by contract CDRN-1306-03961 from the Patient Centered Outcomes Research Institute and UL1TR001433 from the National Center for Advancing Translational Sciences, National Institutes of Health.

## References

1. Coulter A. Engaging patients in healthcare. New York (NY): McGraw-Hill Education; 2011.
2. Carman KL, Dardess P, Maurer M, et al. Patient and family engagement: a framework for understanding the elements and developing interventions and policies. *Health Affairs*. 2013; 32(2): 223–231. [PubMed: 23381514]
3. Concannon TW, Meissner P, Grunbaum JA, et al. A new taxonomy for stakeholder engagement in patient-centered outcomes research. *J Gen Intern Med*. 2012; 27(8):985–991. [PubMed: 22528615]
4. Elwyn G, Crowe S, Fenton M, et al. Identifying and prioritizing uncertainties: patient and clinician engagement in the identification of research questions. *J Eval Clin Pract*. 2010; 16(3):627–63. [PubMed: 20482747]
5. Barello S, Graffigna G, Vegni E, Vegni E. Patient engagement as an emerging challenge for healthcare services: mapping the literature. *Nurs Res Pract*. 2012:59.
6. Greenhalgh T. Chronic illness: beyond the expert patient. *BMJ*. 2009; 338(7695):629–631.
7. Chang BH, Hendricks AM, Slawsky MT, et al. Patient recruitment to a randomized clinical trial of behavioral therapy for chronic heart failure. *BMC Med Res Methodol*. 2004; 4(1):1.
8. Frank G. Current challenges in clinical trial patient recruitment and enrollment. *SoCRA Source*. 2004 Feb.2:30–38.
9. Domecq JP, Prutsky G, Elraiyah T, et al. Patient engagement in research: a systematic review. *BMC Health Serv Res*. 2014; 14(1):89. [PubMed: 24568690]
10. IOM. *Toward Environmental Justice: Research, Education, and Health Policy Needs*. Washington, DC: Institute of Medicine, Committee on Environmental Justice, Health Sciences Policy Program, Health Sciences Section; 1999.
11. IOM. *Dietary Reference Intakes for vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc*. Washington, DC: National Academy Press; 2001.
12. George M, Freedman TG, Norfleet AL, et al. Qualitative research-enhanced understanding of patients' beliefs: results of focus groups with low-income, urban, African American adults with asthma. *J Allergy Clin Immunol*. 2003; 111(5):967–973. [PubMed: 12743559]
13. Saha S, Jacobs EA, Moore RD, et al. Trust in physicians and racial disparities in HIV care. *AIDS patient care and STDs*. 2010; 24(7):415–420. [PubMed: 20578909]
14. Lupton D. The commodification of patient opinion: The digital patient experience economy in the age of big data. *Sociol Health Illn*. 2014; 36(6):856–869. [PubMed: 24443847]

15. Selby JV, Lipstein SH. PCORI at 3 years—progress, lessons, and plans. *N Engl J Med*. 2014; 370(7):592–595. [PubMed: 24521104]
16. Selby JV, Beal AC, Frank L. The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda. *JAMA*. 2012; 307(15):1583–1584. [PubMed: 22511682]
17. Patient Protection and Affordable Care Act, Pub L No. 111–148, 124 Stat 727, §6301.
18. Margolis R, Derr L, Dunn M, et al. The National Institutes of Health' Big Data to Knowledge (BD2K) initiative: capitalizing on biomedical big data. *J Am Med Inform Assoc*. 2014; 21(6):957–958. [PubMed: 25008006]
19. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015; 372(9):793–795. [PubMed: 25635347]
20. Kaushal R, Hripcsak G, Ascheim DD, Bloom T, et al. Changing the research landscape: The New York City Clinical Data Research Network. *J Am Med Inform Assoc*. 2014; 21(4):587–590. [PubMed: 24821739]
21. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006; 3:77–101.
22. Boyatzis R. Transforming qualitative information: Thematic analysis and code development. Thousand Oaks, CA: Sage Publications; 1998.
23. Lincoln YS, Guba EG. *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications; 1985.
24. Goytia CN, Todaro-Rivera L, Brenner B, et al. Community capacity building: a collaborative approach to designing a training and education model. *Prog Community Health Partnersh*. 2013; 7(3):291–299. [PubMed: 24056511]
25. Kim KK, Joseph JG, Ohno-Machado L. Comparison of consumers' views on electronic data sharing for healthcare and research. *J Am Med Inform Assoc*. 2015; 22(4):821–30. [PubMed: 25829461]
26. Ancker JS, Silver M, Miller MC, et al. Consumer experience with and attitudes toward health information technology: a nationwide survey. *Journal of the American Medical Informatics Association*. 2013; 20(1):152–6. [PubMed: 22847306]
27. Mullins CD, Abdulhalim AM, Lavalley DC. Continuous patient engagement in comparative effectiveness research. *JAMA*. 2012; 307(15):1587–1588. [PubMed: 22511684]
28. Gibson B, Butler JM, Lewis L, et al. Predictors of patient's intentions to participate in pragmatic clinical trials: An initial exploration. *SSM-Population Health*. 2016; 2:90–94. [PubMed: 29349131]
29. Willis KF, Robinson A, Wood-Baker R, et al. Participating in research exploring participation and engagement in a study of self-management for people with chronic obstructive pulmonary disease. *Qual Health Res*. 2011; 21(9):1273–1282. [PubMed: 21474655]
30. Bates I, Akoto AY, Ansong D, et al. Evaluating health research capacity building: an evidence-based tool. *PLoS Med*. 2006 Jul 18.3(8):e299. [PubMed: 16942394]
31. Crawford K, Schultz J. Big data and due process: Toward a framework to redress predictive privacy harms. *BCL Rev*. 2014; 55:93.
32. Neff G. Why big data won't cure us. *Big data*. 2013; 1(3):117–123. [PubMed: 25161827]
33. Lohr S. Big Data is opening doors, but maybe too many. Mar 23, 2013. Available at: <http://www.nytimes.com/2013/03/24/technology/big-data-and-a-renewed-debate-over-privacy.html>. Accessed May 18 2016
34. Schadt EE. The changing privacy landscape in the era of big data. *Mol Syst Biol*. 2012; 8(1):612. [PubMed: 22968446]
35. Nieva VF, Sorra J. Safety culture assessment: a tool for improving patient safety in healthcare organizations. *Qual Saf Health Care*. 2003; 12(suppl 2):ii17–ii23. [PubMed: 14645891]
36. Blobel B, Lopez DM, Gonzalez C. Patient privacy and security concerns on big data for personalized medicine. *Health Technol Assess*. 2016:1–7.
37. Scarbrough H, D'Andreta D, Marabelli M, et al. Dominant cognitive frames and the innovative power of social networks. *Acad Manage Rev*. 2016; 2015(1):16147.

**Table 1**

## Pros and Cons of NYC CDRN

<b>PROs of CDRN</b>	
Exposure to research that's going on	P
Give doctors right tools to better patient care	P
Can Inform Policy	P
Can provide more accurate data *	P, C
Can make important information accessible and understandable *	P, C
Can use data to activate community and doctors *	P,C
Tool to organize communities to ask and answer questions *	P
Shed light on diseases, treatments *	P, C
Long term can use to reduce redundancies in testing	C

<b>CONs of CDRN</b>	
Data can never be truly de-identified	P
Don't want people from other hospitals to have their data	P
Worry can't control how many times patients could be contacted	P
Not sure will trust that data is accurate (doubt other datasets like A1c registry) *	P, C
Don't trust info will be shared with communities *	P, C
May share info with insurers *	P, C
Labor intensive *	P, C
Don't see direct benefit for patients *	P, C
IT personnel already overwhelmed at institutions	C
Priority is patient care NOT research	C
Not very interested in the data that will be generated *	P, C
Will never get frontline clinicians to give regular input	C
Concern that data from CDRN will conflict with what doctors recommend to patients	C
Won't help with what they need (i.e. reducing 30d readmission, evaluating treatment effectiveness)	C

\* denotes patient and clinician agreement

**Table 2**

Suggestions for publicizing the NYC-CDRN

<b>PUBLICISING CDRN</b>	
Use local ethnic media to send messages	P
Develop excellent PR campaign, multipronged so people are familiar *	P, C
Share Research 101	P
Give out family/patient/clinician friendly portal to search for research studies *	P, C
Share positive research stories that are relatable *	P, C
Tell people where data will sit, how it will be secured*	P, C
Send clinicians emails with periodic updates about CDRN and trials of potential interest	C

\* denotes patient and clinician agreement

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3**

## Suggestions for Research Projects

<b>RESEARCH SUGGESTIONS</b>	
Topics that match community concerns	P
Comparing treatments*	P, C
Identify and compare neighborhoods with good and bad outcomes and determine related factors*	P, C
Need capacity building to translate perceived problems into research questions*	P, C
Need to have policymakers at table so research also meets their needs	P
Social media for young, face to face for older	P
Hold research forums to generate ideas with patients	C
Use data to create research registry for NYS for providers	C
How to reduce disease transmission in families	C
CF-infertility, MCA, pancreatitis	C
Get ideas through research forums	C
Need to prepare people before meetings so come with ideas*	P, C
Percentage of weight loss with post bariatric surgery compared to age	P
Correlation of bariatric surgeries and attendance a support groups	P
Success rates of surgeries at various institutions	P
Complication rates based on type of surgery and year post surgery	P
Pregnancy complications post-surgery	P
Fertility issues post-surgery (male/female)	P
# of patients post bariatric surgery w/plastic surgery	P
Medications post-surgery	P

\* denotes patient and clinician agreement

**Table 4**

## Possible Recruitment Suggestions and Challenges

<b>RECRUIT SUGGESTIONS</b>	
Doctors as gatekeepers (pros and cons about this) *	P, C
CHW's/coordinators to introduce research and engage patients *	P, C
Social service, clergy, advocates as gatekeepers	P
Go through trusted organizations	P
Transparency with possible benefits to patients *	P, C
Identify and honor language preferences	P
No texts-don't trust from stranger (could use for follow-up)	P
Text and email patients	C
Mail or phone call from someone from their institution	P
Approach person in waiting room if recruiter from that site.	P
Don't cold contact patients	C
Use fundraisers for recruitment	P
Use local research champions within practices	C
Have dedicated research staff to speak/enroll patient (warm handoff)	C
Research staff should look like patients	P
New consent strategies (video consent)	C
Social media to advertise for research (especially for young) *	P, C
Create referral tool and process for docs to use	C
Incentives for patients	C
<b>RECRUIT CHALLENGES</b>	
No clinical/research separation (baby/bathwater)	P
Research futilism (research to date hasn't made impact)	P
Trust-Petri dish/victimization concerns	P
Acknowledge past abuses in research before taking part	P
Don't trust people outside their own hospital/clinic	P
Don't direct contact-will scare people that their info was released	P
Want to know who sponsors research before take part *	P, C
Socio-Economic patient issues (several comorbidities)	C
Consenting when patients have an appointment is too time consuming	C
Consenting by EMR is too time consuming for provider	C
Patients scared of term "research"	C
Disruption to clinical workflow	C
Time burden on patient	P

\* denotes patient and clinician agreement